

# Shift Invariant Ear Feature Extraction using Dual Tree Complex Wavelet Transform for Ear Recognition

Rajesh M Bodade, Maj Jayesh Nayyar  
 Military College of Telecommunication Engineering  
 Mhow, Indore, India  
 rajeshbodade@gmail.com  
 Indian Army, India  
 jayeshnayyar@gmail.com

**Abstract**— Since last 10 years various methods have been used for ear recognition. This paper describes the automatic localization of an ear and its segmentation from the side pose of face image. In this paper, authors have proposed a novel approach of feature extraction of iris image using 2D Dual Tree Complex Wavelet Transform (2D-DT-CWT) which provides six sub-bands in 06 different orientations, against 3 orientations in DWT. DT-CWT being complex it exhibit the property of shift invariance. Ear features vectors are obtained by computing mean, standard deviation, energy and entropy of these six sub-bands DT-CWT and three sub-bands of DWT Canberra distance and Euclidian distance are used for matching. The accuracy of recognition is achieved above 97 %.

**Keywords**- Ear recognition; ear detection; ear biometrics; DT-CWT; complex wavelet transform

## I. INTRODUCTION

Ear recognition has received considerably less attention than many alternative biometrics, including face, fingerprint and iris recognition. Ear-based recognition is of particular interest because it is non-invasive, and because it is not affected by environmental factors such as mood, health, and clothing. Also, the appearance of the auricle (outer ear) is relatively unaffected by aging, making it better suited for long-term identification [1]. Ear images can be easily taken from a distance without knowledge of the person concerned. Therefore ear biometric is suitable of surveillance, security, access control and monitoring applications.

As compared to face biometrics [2]-[4] ears have several advantages over complete faces, like, reduced spatial resolution, a more uniform distribution of color, and less variability with expressions and orientation of the face. Its deep three-dimensional structure makes it very difficult to counterfeit. In face recognition there can be problems of illumination variation, pose variation and facial expressions [4]. Ear was first used for recognition of human being by Iannarelli who used manual techniques to identify ear images. The medical literature provides information that ear growth is proportional after first four months of birth and changes are not noticeable in the age 8 to 70 [1]. The remainder of this paper consists of: existing ear recognition techniques, localisation and normalisation of ear, feature extraction using DT-CWT, matching, experimental results and conclusions covered in Section-2, 3, 4, 5, 6 and 7 respectively.

## II. EXISTING EAR RECOGNITION TECHNIQUES

Major work on automatic ear localisation [2]-[11] has been done recently in past 10 years. Automatic ear recognition using Voronoi diagrams to take care of adverse effects of lighting, shadowing and occlusion has been presented by Burge and Burger [6]. In [9], Active Contour Model (or Snakes) is used to segment the ear from the side images of the face. Hurley, Nixon and Carter [7] have used force field transformations for ear localisation. [8][12] and [13] make use of 3-D range images to extract the ear from the image of a human. However, the tougher challenge is to detect the ear from an intensity image. A shape model-based technique for locating human ears in side face range images is proposed in [8]. In this method, the ear shape model is represented by a set of discrete 3D vertices corresponding to ear helix and anti-helix parts. Ansari and Gupta have proposed an approach based on edges of outer ear helices by exploiting the parallelism between the outer helix curves of the ear to localize the ear [10]. Skin-color and contour information has been exploited for ear detection by Yuan and Mu [11]. In [13], authors have presented a distance transform and template based technique for automatic ear localization from a side face image. The technique first segments skin and non-skin regions in the face and then uses template based approach to find the ear location within the skin regions.

Victor et al. [4] and Chang et al. [2] have researched use of PCA and FETET for ear recognition. Moreno et al. [5] used 2D intensity images of ears with three neural net approaches for ear recognition. In [16], Anupama Sana et al. presented an ear biometric system based on discrete Haar Wavelet Transform whereas Wang and Yuan [17] used Gabor wavelets and general discernment analysis. Wang Xiaoyun et al. [19] proposed block segmentation based approach whereas modular neural network architecture has been proposed by Gutierrez et al. [18].

## III. PROPOSED SYSTEM

The block diagram of proposed system is shown in Fig 1.



Figure 1. Block diagram of proposed system

It consist of image acquisition module, preprocessing and automatic ear localization module, DT-CWT based feature extraction module and matching module.

#### A. IMAGE ACQUISITION MODULE

Sony DSC-HX1 (15 Megapixel and optical zoom of 20x) camera is used for image acquisition. Database of 240 images of 40 subjects for left and right ears is crated at MCTE and UND database is also used.

#### B. PREPROCESSING AND EAR LOCALISATION MODULE

The raw image is not suitable for feature extraction due to its large background thus some pre-processing is required to make it suitable.

#### C. FEATURE EXTRACTION MODULE

After successful ear localization, features are extracted using DT-CWT. The details of it is stated in section IV.

#### D. MATCHING MODULE

Energy, Entropy, Mean and Standard Deviation of each sub-bands of DT-CWT is calculated to create a feature vector. Euclidian distance and Canberra distance are used as similarity measure for matching the feature vectors of test image with that of images stored in database(1:N) match.

### IV. AUTOMATIC EAR LOCALISATION

This was a very challenging task as most of the work carried out on this aspect is in experimental stage. The algorithm so designed includes the finer points of various algorithms and additional measures to try and further enhance and improve the ear localization results. The algorithm works as under.

- (i) Take a side face image of an individual (under varied background and lighting conditions).
- (ii) Since *RGB* representation of color images is not suitable for characterizing skin-color, it first converts the *RGB* color space to chromatic color space, and then uses the chromatic color information for further processing.
- (iii) Apply threshold to the resultant image and convert to binary.
- (iv) A few results under varied background conditions are shown in Figure 2 to Figure 4.



Figure 2. Side face image with open sky background



Figure 3. Side face image with class room background



Figure 4. Side face image with class room background from UND database

(v) Once the background is removed from the image we determine the nose tip of the subject and assuming an approximate distance between the subject's nose tip and ear pit and the average size of human ear we crop the ear part from the side face of the image.

(vi) Thereafter, we take a four pronged approach to determine the edge variations from top (at Helix), from side (at Concha and at Helix posterior) and from bottom (at Lobulo) to determine the ear edges and thus crop out the exact ear out of the image.

(vii) A two resulting ear images localized automatically from the side face images are shown in Figure 5.

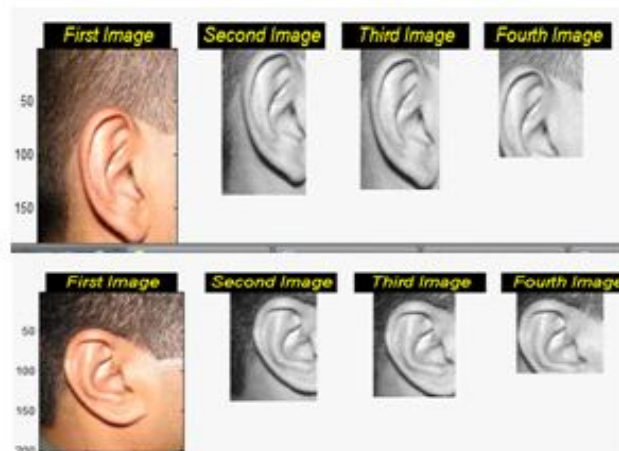


Figure 5. Automatic cropping for Ear localization

(viii) The cropped ear images may be of varying sizes so the feature set of images may also vary. Hence the images are normalized to a constant size.

### V. FEATURE EXTRACTION

DT-CWT is formulated by Kingbury and Selesnick[20], [21] using two trees (real and imaginary tree) of DWTs with different filter real coefficients for imaginary tree filters designed from the coefficients of real tree filters to overcome the limitations of DWTs. The details of DT-CWT and feature extraction are stated in following subsections.

#### A. DUAL TREE COMPLEX WAVELET TRANSFORM

In dual-tree, two real wavelet trees are used as shown in Figure 6, each capable of perfect reconstruction (PR). One tree generates the real part of the transform and the other is used in generating complex part[20]. As shown,  $\{H_0(z), H_1(z)\}$  is a Quadrature Mirror Filter (QMF) pair in the real-coefficient analysis branch. For the complex part,  $\{G_0(z), G_1(z)\}$  is another QMF pair in the analysis branch. All filter pairs are orthogonal and real-valued.

It has been shown [21] that if filters in both trees be made to be offset by half-sample, two wavelets satisfy Hilbert transform pair condition and an approximately analytic wavelet is given by Eq(1).

$$\psi(x) = \psi_h(x) + j\psi_g(x) \quad (1)$$

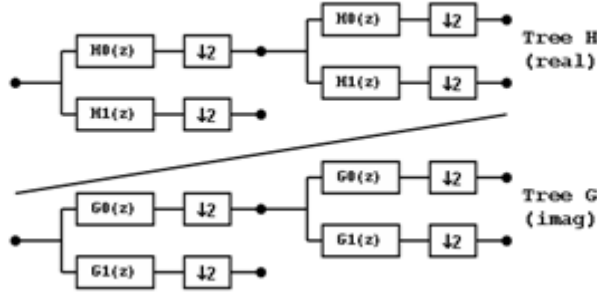


Figure 6. Selesnick's Dual Tree DWT

Thus, if  $G_0(\omega) \cong H_0(\omega) \times e^{-i\theta(\omega)}$  and  $\theta(\omega) \cong \omega/2$

then  $\psi_g(\omega) \cong -j\psi_h(\omega), \omega > 0$

$$\cong j\psi_h(\omega), \omega < 0 \quad (2)$$

From Eq(1) and (2), low pass filters after the first stage and at first stage respectively are given by Eq(3):

$$g_0(n) = h_0(n-0.5) \text{ and } g_0(n) = h_0(n-1) \quad (3)$$

Similar relations also hold true for high pass filters of both the trees..

In this algorithm, (10,10)-Tap near orthogonal wavelet filters are used in first stage and 'db7' filters are used for higher stages in the real tree (i.e.  $h_0$  and  $h_1$ ) [20]. The imaginary low pass filter is derived from the above half sample delayed condition.

The high pass filter is the quadrature-mirror filter of the low pass filter. The reconstruction filters are obtained by time reversal of decomposition filters. All the filters used are of same length based on Selesnick's approach [20], [21], [23], [24] unlike Kingsbury's approach.

The 2D separable DWT can be written in terms of 1D scaling functions ( $\phi$ ) and wavelet functions ( $\psi$ ) as:

$$\begin{aligned} \psi^0(x, y) &= \phi(x)\psi(y) \\ \psi^{90}(x, y) &= \psi(x)\phi(y) \\ \psi^{\pm 45}(x, y) &= \psi(x)\psi(y) \end{aligned} \quad (4)$$

Oriented non-separable 2D wavelet transform is derived by combining the sub-bands of two separable 2D DWTs. The pair of conjugate filters are applied to two dimensions (x and y), which can be expressed by Eq(5) as given bellow:

$$(h_x + jg_x)(h_y + jg_y) = (h_x h_y - g_x g_y) + j(h_x g_y + h_y g_x) \quad (5)$$

The filter bank structure of 2D DT CWT, to implement Eq(5) is shown in Figure 7.

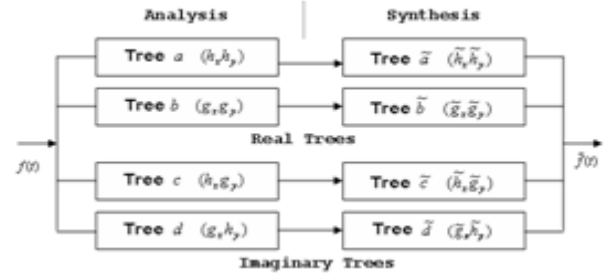


Figure 7. Filter bank structure of 2D DT CWT

Tree-a and Tree-b is combined to compute the Real part of Eq(5) i.e Real (2D DWT) tree of CWT as shown in Figure 8. Similarly, Imaginary (2D DWT) tree of CWT can be obtained from tree-c and tree-d i.e.  $(h_x g_y - g_x h_y)$ , as per Eq(5).

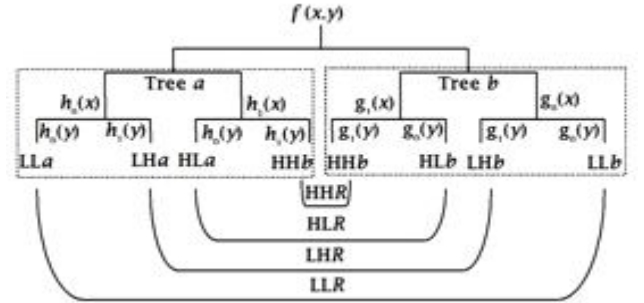


Figure 8. Formation of Real Tree DT CWT

Thus, the decomposition for each mode is performed in a standalone mode, in one after another stage i.e. total of 6 detailed coefficients are derived at each stage; three for the real tree and three for the imaginary tree. 3-stage decomposition is performed. At each stage, coefficients are oriented towards their respective directions as stated in Eq(4). Following six wavelets, as given by Eq(6), are used to obtain oriented 2-D separable wavelets [20]:

$$\begin{aligned} \psi_{1,1}(x, y) &= \phi_h(x)\psi_h(y), & \psi_{2,1}(x, y) &= \phi_g(x)\psi_g(y), \\ \psi_{1,2}(x, y) &= \psi_h(x)\phi_h(y), & \psi_{2,2}(x, y) &= \psi_g(x)\phi_g(y), \\ \psi_{1,3}(x, y) &= \psi_h(x)\psi_h(y), & \psi_{2,3}(x, y) &= \psi_g(x)\psi_g(y), \end{aligned} \quad (6)$$

where,  $\psi_{1,i}$  correspond to the coefficients derived from the real tree and  $\psi_{2,i}$  correspond to the coefficients derived from the imaginary tree. They can be combined by Eq(7) to form complex wavelet coefficients.

$$\begin{aligned} \psi_i(x, y) &= \frac{1}{\sqrt{2}}(\psi_{1,i}(x, y) - \psi_{2,i}(x, y)), \\ \psi_{i+2}(x, y) &= \frac{1}{\sqrt{2}}(\psi_{1,i}(x, y) + \psi_{2,i}(x, y)) \end{aligned} \quad (7)$$

Normalization by  $1/\sqrt{2}$  is used so that the sum difference operation constitutes an ortho-normality. These six wavelet sub-bands of the 2-D DT-CWT are strongly oriented in  $\{+15^\circ, +45^\circ, +75^\circ, -15^\circ, -45^\circ, -75^\circ\}$  direction as shown in fig(5) by red lines and it captures image information in those directions. Thus, in particular, 2D dual-tree wavelets are not only approximately analytic but also oriented and are shift invariant because of its analytic structure [20]. The impulse responses of three 2-D sub-bands (2-D non separable filters for detailed coefficients)

of DWT and six sub-bands (2-D non separable filters for detailed coefficients) of DT-CWT are shown in Figure 9.

### B. FEATURE EXTRACTION

Ear analysis using DWT provides singularities (edges) in only three directions (0, 45, 90) and without phase information which is improved by finding the singularities, with phase information, in six directions (0, +/-15, +/-30, +/-45, +/-60, +/-75, 90) and at many freq bands using DT-CWT to achieve shift invariant features for better accuracy and efficiency at less computational cost as compared to existing methods.

From the detailed study of prevalent techniques already employed for ear recognition, it is realized that nobody had made use of Complex Wavelets for ear recognition. This realization laid the foundation of utilizing this approach to determine whether or not the said approach can further enhance and improve the recognition rates already achieved by other methods. As the advantages of using CWT vis-à-vis DWT, it is imperative to employ Dual Tree - Complex Wavelet Transform (Selesnick) (DT-CWT(S)) for this work.

The DT-CWT(S) algorithm is used to design and implement the Dual Tree structure (up to Level 2) using MATLAB, employing first stage and second stages low pass and high pass filter coefficients given by Selesnick. The impulse responses of three 2-D sub-bands (2-D non separable filters for detailed coefficients) of DWT and six sub-bands (2-D non separable filters for detailed coefficients) of DT-CWT are shown in Figure 9.

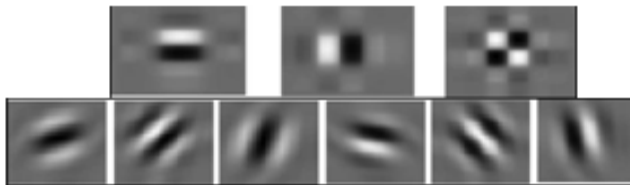


Figure 9. Impulse responses of sub-bands of DWT and DT-CWT.

DT-CWT has 06 directional wavelets oriented at angles of  $\pm 15, \pm 45, \pm 75$  in 2-Dimension. We get these six directions by passing the 2-D signal (image) through the real tree structure using the filter coefficients of both real and imaginary trees. The wavelet coefficients of each image which formed part of the Training Database were thus obtained in all the six directions and stored for further matching and testing. These directions can be seen clearly from the Figure 10 which represents the Level 1 decomposition of an image.

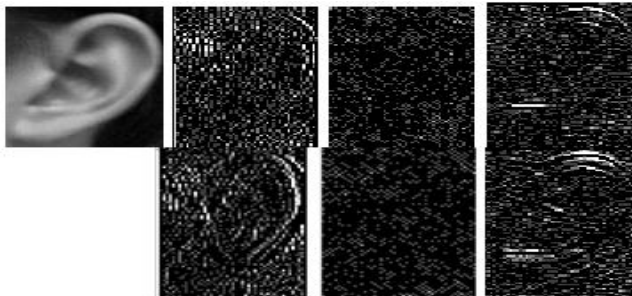


Figure 10. Real and imaginary tree wavelet sub-band images

## VI. EXPERIMENTAL RESULTS

All the Training images of both the databases (MCTE database of 240 images of 40 subjects for right and left ears and UND database of 219 subjects under J-collection and G-collection) are processed and their respective wavelet coefficients at Level 1 and Level 2 are calculated. Energy, Entropy, Mean and Standard Deviation of each image's wavelet coefficient are then calculated and stored in an MS Access database. Thereafter images from the Test Set and random images were matched with these stored values using Euclidean and Canberra distance matching techniques and results for False Acceptance Rate (FAR), False Rejection Rate (FRR), Equal Error Rate (EER) and Receiver's Operating Curve (ROC) compiled at various thresholds. All the results are stated in Table 1.

Figure 11 to 14 shows the FAR, FRR and ROC of best and worst case of Canberra distance and best and worst case of Euclidean distance when tested on database with following details.

Name Of Data-Base : UND – Collection G

No of Images in the Training database: 20

No of Images in Test Database: 90

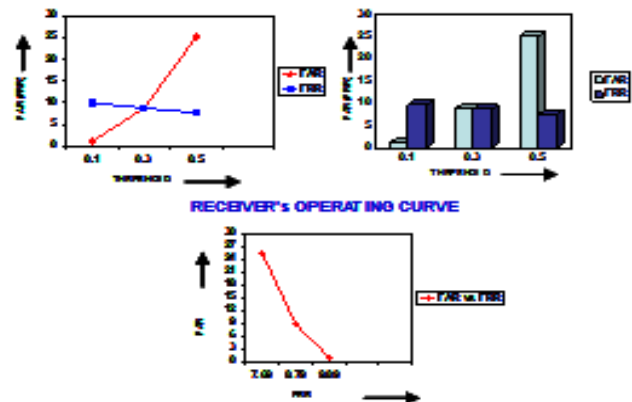


Figure 11. Results using Canberra distance and feature vector of energy only (Worst case of Canberra Distance)

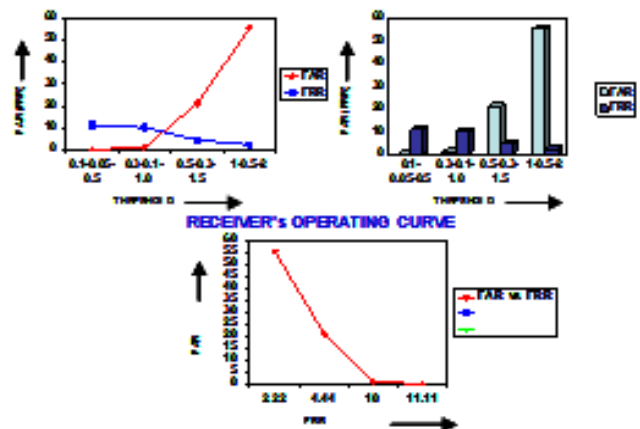


Figure 12. Results using Canberra distance and feature vector of energy + std deviation + entropy (Best case of Canberra Distance)



TABLE I. COMPILED RESULTS OF AVG FAR, AVG FRR AND AVG RECOGNITION RATE USING CANBERRA AND EUCLIDIAN DISTANCE FOR DIFFERENT FEATURE VECTORS

Feature Vector / Distance	Canberra Distance			Euclidian Distance		
	Avg FAR	Avg FRR	Recognition Rate	Avg FAR	Avg FRR	Recognition Rate
Energy only	8.79 %	8.79 %	91.21 %	16.66 %	5.55 %	83.34 %
Entropy only	6.59 %	8.79 %	93.41 %	12.22 %	6.66 %	87.88 %
Standard Deviation only	8.79 %	8.79 %	91.21 %	21.11 %	4.44 %	78.89 %
Energy+Entropy	3.29 %	8.79 %	96.71 %	3.33 %	8.88 %	96.67 %
Energy+Standard Deviation	5.49 %	4.39 %	94.51 %	4.44 %	6.66 %	95.56 %
Entropy+Standard Deviation	5.49 %	8.79 %	94.51 %	6.66 %	4.44 %	93.34 %
Energy + Entropy + Standard Deviation	2.19 %	7.69 %	97.81 %	3.33 %	7.77 %	96.67 %

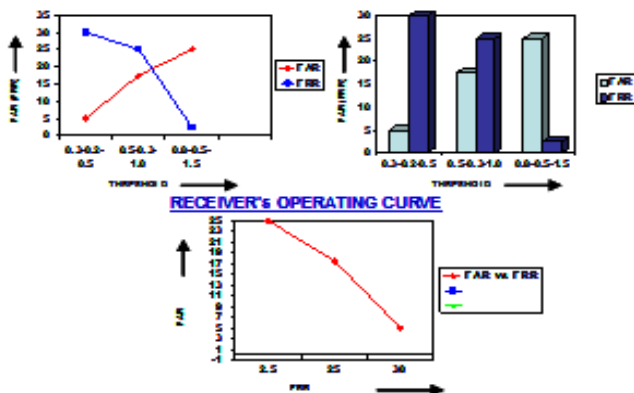


Figure 13. Results using Euclidian distance and feature vector of Std Deviation only (Worst case of Euclidian Distance)

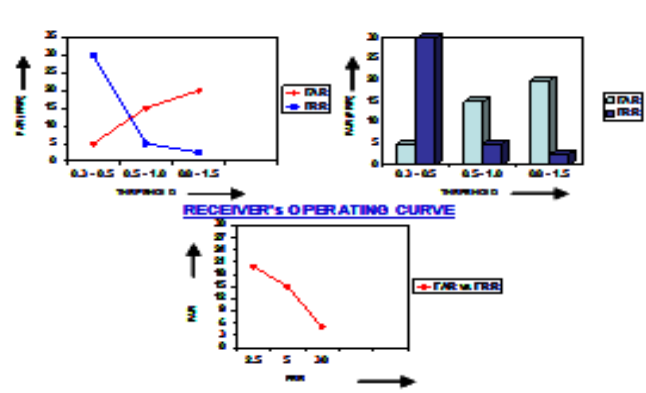


Figure 15. Results of DWT using Canberra distance and feature vector of energy + std deviation + entropy (Best case of DWT)

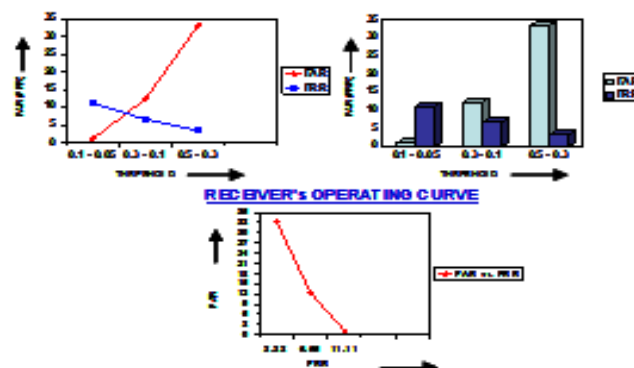


Figure 14. Results using Euclidian distance and feature vector of energy + std deviation + entropy (Best case of Euclidian Distance)

The maximum recognition rate of 81% is obtained when DWT is used for feature extraction and Canberra distance is used as similarity metric for combined vector of energy, std. deviation and entropy. FAR, FRR and ROC for it is shown in figure 15.

## CONCLUSIONS

The authors have introduced a new 2D DT CWT for ear recognition first time because of its ability to capture shift invariant features in 06 orientations. The experimental results have demonstrated the effectiveness of the proposed method in terms of improving the recognition rate.

Canberra distance has shown better results than Euclidian distance because it normalises the individual feature components before finding the distance between the two images.

The best recognition rate of over 97% has been achieved using Canberra distance when feature vectors of energies, standard deviation and entropy of sub-bands of DT-CWT are used together.

The authors are working on improving the recognition rate by using the RCWF [22], [23], [24] in combination with DT-CWT to obtain features in 12 orientations (06 by DT-CWT and 06 by RCWF).

# REFERENCES

- [1] Alfred Iannarelli, "Ear Identification," Forensic Identification Series, Paramount Publishing Company, Fremont, California, 1989.
- [2] K. Chang., K.W. Bowyer, S. Sarkar, and B. Victor, "Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 9, pp. 1160-1165, September 2003.
- [3] D.J. Hurley, M.S. Nixon, J.N. Carter, "A New Force Field Transform for Ear and Face Recognition," Proc. IEEE 2000 International Conference on Image Processing (ICIP 2000), IEEE Press, pp. 25-28.
- [4] B. Victor, K.W. Bowyer, S. Sarkar, "An evaluation of face and ear biometrics" Proc. International Conference on Pattern Recognition, August 2002, pp. 429-432.
- [5] B. Moreno, A. Sánchez, J.F. Vélez., "On the Use of Outer Ear Images for Personal Identification in Security Applications," Proc. IEEE 33 Annual International Carnahan Conference on Security Technology, 1999, pp. 469-476.
- [6] M. Burge and W. Burger, "Ear biometrics in computer vision," Proc. International Conference on Pattern Recognition, (ICPR'00), vol. 02, 2000, pp. 822-826.
- [7] D. J. Hurley, M. S. Nixon, and J. N. Carter, "Automatic ear recognition by force field transformations," Proc. IEE Colloquium: Visual Biometrics, 2000, pp. 8/1-8/5.
- [8] H. Chen and B. Bhanu, "Human ear detection from side face range images," Proc. International Conference on Pattern Recognition, (ICPR'04), vol. 3, IEEE Computer Society Press, 2004, pp. 574-577.
- [9] L. Alvarez, E. Gonzalez, and L. Mazorra, "Fitting ear contour using an ovoid model," in IEEE International Carnahan Conference on Security Technology, (ICCST'05), 2005, pp. 145-148.
- [10] S. Ansari and P. Gupta, "Localization of ear using outer helix curve of the ear," Proc. (International Conference on Computing: Theory and Applications, ICCTA'07), IEEE Computer Society Press, 2007, pp. 688-692.
- [11] L. Yuan and Z.-C. Mu, "Ear detection based on skin-color and contour information," Proc. International Conference on Machine Learning and Cybernetics, (ICMLC'07), vol. 4, (Hong-Kong), 2007, pp. 2213-2217.
- [12] Hui Chen and Bir Bhanu, "Human Ear Recognition in 3D", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol 29, No 4, April 2007
- [13] Surya Prakash, Umarani Jayaraman and Phalguni Gupta, "Ear Localization from Side Face Images using Distance Transform and Template Matching", Proc. IEEE Int'l Workshop on Image Processing Theory, Tools and Applications, (IPTA 2008), Sousse, Tunisia, Nov 2008 pp. 1-8.
- [14] D. Hurley, M. Nixon, and J. Carter, "Force Field Energy Functionals for Ear Biometrics" Computer Vision and Image Understanding, vol.98, 2005, pp.491-512.
- [15] Ping Yan and Kevin W. Bowyer, "Biometric Recognition Using 3D Ear Shape" IEEE Transactions on pattern analysis and machine intelligence, vol. 29, no. 8, August 2007.
- [16] A. Sana, F. Gupta and R. Prukait, "Ear Biometrics: A new approach," Proc. Sixth International Conference on Advances in Pattern Recognition, Jan 2007, pp.46-50.
- [17] Xiaoyun Wang and Weiqi Yuan, "Gabor wavelets and General Discriminant analysis for ear recognition," Proc. 8th World Congress on Intelligent Control and Automation (WCICA 2010), IEEE Press, 2010, pp. 6305 - 6308
- [18] L. Gutierrez, P. Melin, and M. Lopez, "Modular neural network integrator for human recognition from ear images," Proc. International Joint Conference on Neural Networks (IJCNN 2010), IEEE Press, 2010, pp. 1-5.
- [19] Wang Xiaoyun and Yuan Weiqi, "Human ear recognition based on block segmentation," Proc. International Conf on Machine Learning and Cybernetics, 2009, vol. 4, pp. 2418 - 2422.
- [20] Ivan W. Selesnick, Richard G. Baraniuk, and Nick G. Kingsbury, "The Dual Tree Complex Wavelet Transform: A Coherent Framework for Multiscale Signal and Image Processing", IEEE Signal Processing Magazine, Nov 2005, pp. 123-151.
- [21] N.G. Kingsbury, "The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters," in Proc. 8th IEEE DSP Workshop, Utah, Aug. 9-12, 1998, paper no. 86.20.
- [22] M. Kokare, P.K. Biswas, and B.N. Chatterji, "Rotation invariant texture features using rotated complex wavelet for content based image retrieval," in Proc. IEEE Int. Conf. Image Processing, Singapore, Oct. 2004, vol. 1, pp. 393-396.
- [23] Rajesh Bodade and Sanjay Talbar, "Iris Recognition Using Multi-Directional Wavelets: A Novel Approach" Journal of Advances in Engineering Sciences Sect. C (3), July-September 2008, (special issue on Image Processing) ISSN: 0973-9041
- [24] R. Bodade and S. Talbar, "Iris Recognition using Combination of Dual Tree Rotated Complex Wavelet and Dual Tree Complex Wavelet" Proc. IEEE International Conference on Communication-2009 (ICC-2009), IEEE Press, June 2009.
- [25] J. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence," IEEE Trans. PAMI, vol. 15, no. 11, Nov. 1993, pp. 1148-1161.